Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 9023719, 8 pages
https://doi.org/10.1155/2022/9023719

Research Article

Intrusion Detection System for Industrial Internet of Things Based on Deep Reinforcement Learning

Sumegh Tharewal,[1] **Mohammed Waseem Ashfaque**,[2] **Sayyada Sara Banu**,[3] **Perumal Uma**,[4] **Samar Mansour Hassen**,[5] **and Mohammad Shabaz**[6]

[1]*School of Computer Science, Dr. Vishwanath Karad MIT World Peace University, Kothrud, Pune, India*
[2]*Department of IT, Al Buraimi University College, Buraimi, Oman*
[3]*Jazan University, Department of CS & IT, Saudi Arabia*
[4]*Department of CS & IT, Jazan University, Saudi Arabia*
[5]*Department of CS & IT, MBA, Jazan University, Saudi Arabia*
[6]*Arba Minch University, Ethiopia*

Correspondence should be addressed to Sumegh Tharewal; sumeghtharewal@gmail.com,
Mohammed Waseem Ashfaque; waseem2000@gmail.com, and Mohammad Shabaz; mohammad.shabaz@amu.edu.et

Received 26 January 2022; Revised 11 February 2022; Accepted 15 February 2022; Published 7 March 2022

Academic Editor: Deepak Kumar Jain

The Industrial Internet of Things has grown significantly in recent years. While implementing industrial digitalization, automation, and intelligence introduced a slew of cyber risks, the complex and varied industrial Internet of Things environment provided a new attack surface for network attackers. As a result, conventional intrusion detection technology cannot satisfy the network threat discovery requirements in today's Industrial Internet of Things environment. In this research, the authors have used reinforcement learning rather than supervised and unsupervised learning, because it could very well improve the decision-making ability of the learning process by integrating abstract thinking of complete understanding, using deep knowledge to perform simple and nonlinear transformations of large-scale original input data into higher-level abstract expressions, and using learning algorithm or learning based on feedback signals, in the lack of guiding knowledge, which is based on the trial-and-error learning model, from the interaction with the environment to find the best good solution. In this respect, this article presents a near-end strategy optimization method for the Industrial Internet of Things intrusion detection system based on a deep reinforcement learning algorithm. This method combines deep learning's observation capability with reinforcement learning's decision-making capability to enable efficient detection of different kinds of cyberassaults on the Industrial Internet of Things. In this manuscript, the DRL-IDS intrusion detection system is built on a feature selection method based on LightGBM, which efficiently selects the most attractive feature set from industrial Internet of Things data; when paired with deep learning algorithms, it effectively detects intrusions. To begin, the application is based on GBM's feature selection algorithm, which extracts the most compelling feature set from Industrial Internet of Things data; then, in conjunction with the deep learning algorithm, the hidden layer of the multilayer perception network is used as the shared network structure for the value network and strategic network in the PPO2 algorithm; and finally, the intrusion detection model is constructed using the PPO2 algorithm and ReLU (R). Numerous tests conducted on a publicly available data set of the Industrial Internet of Things demonstrate that the suggested intrusion detection system detects 99 percent of different kinds of network assaults on the Industrial Internet of Things. Additionally, the accuracy rate is 0.9%. The accuracy, precision, recall rate, F1 score, and other performance indicators are superior to those of the existing intrusion detection system, which is based on deep learning models such as LSTM, CNN, and RNN, as well as deep reinforcement learning models such as DDQN and DQN.

1. Introduction

Industrial Internet of Things is a term that refers to the use of Internet of Things technologies to the industrial sector. Its primary objective is to integrate and advance industrial automation and Internet of Things technologies. The Industrial Internet of Things enabled a previously unheard-of integration of production, monitoring, and management subsystems. The control center houses a variety of systems. Unified management enables more effective processing of diverse industrial data. Due to its complexity and openness, the Industrial Internet of Things faces increasing network security threats. The National Internet Emergency Center CNCERT released a situation review in 2019 indicating that about 41% of Industrial Internet of Things devices had high-risk vulnerabilities and hidden hazards. Monitoring revealed that 2249 sets of networked monitoring and management systems were exposed in key sectors such as electricity, oil and gas, and urban rail transit, with 653 sets of electricity, 584 sets of oil and natural gas, and 100 sets of urban rail transit [1]. Industrial IoT is a complicated network. Any failure or irregularity in a component of the system may quickly cause catastrophic harm to the whole system. As a result, detecting network assaults quickly and effectively is critical for a fast and effective network response. An intrusion detection system (IDS) is a critical component of network security protection because it enables the system to detect network intrusions efficiently. However, in recent years, as the operating environment and structure of the Industrial Internet of Things have changed, traditional intrusion detection models (such as intrusion detection models based on simple machine learning) have been unable to provide adaptive detection, response, and defence against complex network attacks. The deep reinforcement learning (DRL) method is capable of solving industrial objects efficiently. Uncertainty and other issues in a networked environment were investigated using the agent as the carrier of reinforcement learning. The agent utilizes unknown area and integrates its own experience to learn [2, 3]. Deep reinforcement learning will enhance the learning process's decision-making capacity and combine the perceptual skills of deep understanding, using deep knowledge to conduct simple and nonlinear transformations of large-scale original input data into higher-level abstract expressions and utilizing reinforcement learning or learning based on feedback signals, and, in the absence of guiding information, it is based on the trial-and-error learning model, from the interaction process with the environment to discover the best feasible solution [4, 5]. To address the aforementioned issues, this article presents a novel depth-based reinforcement method, PPO2's intrusion detection system DRL-IDS, for use in the Industrial Internet of Things.

The system extracts the most appealing feature set using a feature selection method based on LightGBM, significantly reducing the computational complexity of the model; it includes three layers of hidden data. The Tibetan layer's multilayer perception serves as the common deep neural network structure for the value network and policy network in the intrusion detection system, thus constructing an intrusion detection system based on the deep reinforcement learning PPO2 algorithm. Finally, the system makes advantage of the ReLU function to minimize overfitting. Rectified Linear Unit (ReLU) is a term that refers to a linear unit that has been recalibrated; the fundamental benefit of the ReLU effect over other activation functions is that this does not simultaneously stimulate all or most of the neurons. The disappearing gradation issue is fixed by using a rectified linear activation function, which allows systems to learn quicker and perform much better. As a classification output, the intrusion detection system described in this article was evaluated against a publicly available data set of the Industrial Internet of Things from the United States Department of Energy's Oak Ridge National Laboratory. The findings indicate that the intrusion detection system is effective. He defeated 99.9% of different kinds of cyberassaults. Additionally, the accuracy rate of 0.9 percent, as well as the accuracy, precision, recall rate, F1 score, and other indicators, outperforms current long-term, short-term memory networks (LSTM); convolution neural networks (CNN); recurrent neural networks (RNN); and other deep learning models, as well as deep double Q networks (DDQN), deep Q networks (DQN), and other deep reinforcement learning model intrusion detection systems.

2. Related Work

Intrusion detection systems are extensively utilised in both conventional industrial control systems and contemporary Industrial IoT [6]. The Industrial Internet of Things is a complex network. Any mistake or inconsistency in a system component might swiftly result in catastrophic damage to the entire system. As a result, swiftly and efficiently identifying network assaults is crucial for an immediate and efficient network response. Because that allows the system to identify network breaches effectively, an intrusion detection system (IDS) is a crucial component of network security protection. In 2019, Jiang et al. [7] presented a deep learning-based IICS anomaly detection system. TCP/IP data packets include information that may be utilised for learning and verification. Sunny et al. [8] utilised Bi-LSTMGRNN to identify Industrial IoT threats in 2020 and trained multilayer deep neural networks using the new UNSWNB15 data set. In 2020, Chen et al. [9] developed a network intrusion detection system based on a CNN-based data collection and monitoring control system to safeguard the Industrial Internet of Things from DDoS attacks and other cyberattacks, cyberassaults against SCADA systems in general and cyberattacks on SCADA systems in particular. Patil and Sankpal [10] investigated power theft attacks in intelligent grids in 2019 and developed a deep learning-based intrusion detection system to identify such cyberassaults. In 2021, Yuan et al. [11] presented the Deep Fed federated deep learning method for detecting and mitigating cyber risks to the dispersed Industrial Internet of Things. However, the aforementioned techniques cannot keep up with

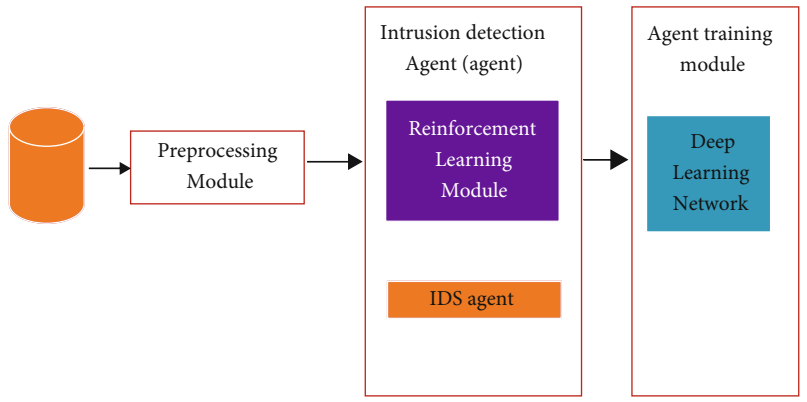

Figure 1: Proposed reinforcement learning based IDS for IOT data.

Table 1: Data set description.

Attack type	Quantity	Attack type description
Normal	l 61156	Normal network traffic
NMRI	2763	Simple malicious response injection
CMRI	15466	Complex malicious response injection
MSCI	78	Malicious status command injection
MPCI	7637	Malicious parameter command injection
MFCI	573	Malicious function command injection
DoS	1837	Denial of service attack
Reconnaissance	6805	Reconnaissance attack

Table 2: PPO2 hyperparameter table.

Hyperparameter name	Meaning	Value
n_env	The number of environment copies running in parallel	15 int
n_steps	The number of steps to run for each environment update	512 int
ent_coef	Loss entropy coefficient	0.000001 float
Learning_rate	Learning rate (can be a function)	Linear schedule from 0.0021 to 0. 0
Max_grad_norm	The maximum value of gradient shear	0.8 float
Minibatches	The number of training batches for each update	16 int

today's high-speed, large-capacity, and complicated multidimensional data. When it comes to Industrial IoT data, a lengthy training procedure is often needed. As a consequence, precision must be increased. The Industrial IoT made it possible to integrate previously unimaginable production, monitoring, and control subsystems. A number of systems are housed in the command center. Integrated management makes it easier to process a wide range of industrial data. The Industrial Internet of Things confronts escalating network security concerns due to the complex nature and openness. Deep reinforcement learning can maximise reward in a known network environment and includes an exploration function that automatically mines more useful information from the network environment, and the model converges fast. Suwannalai and Polprasert [12] presented an intrusion detection system based on reinforcement learning for monitoring and analysing sensor networks in 2020 and compared it to an adaptive machine learning-based intrusion detection system and a cluster hybrid intrusion detection system. Zhou et al. [13] presented a context-adaptive intrusion detection system in 2020 that leverages several independent deep reinforcement learning agents deployed throughout the network to improve the detection accuracy of emerging complex network assaults. Wu and Feng [14] presented a partly observable Markov decision process based on a model-free reinforcement learning method for detecting online network attacks in 2017. Wang et al. [15] presented an adversarial multiagent reinforcement learning model for intrusion detection systems in 2020. However, deep reinforcement learning-based intrusion detection systems still have potential for development and optimization in terms of training efficiency and accuracy. In 2020, Ashiquzzaman et al. [5] presented a deep reinforcement

Table 3: Classification report of DRL-IDS.

Attack type	F1 score	Recall	Precision	Accuracy
Normal	0.9930	0.9924	0.9936	0.9909
NMRI	0.9512	0.9457	0.9568	0.9909
CMRI	0.9943	0.9987	0.9989	0.9909
MSCI	0.9647	0.9588	0.9706	0.9909
MPCI	0.9767	0.9793	0.9741	0.9909
MFCI	0.9737	0.9487	1.0000	0.9909
DoS	0.9855	0.9755	0.9957	0.9909
Reconnaissance	1,000,000	1,000,000	1,000,000	0.9909

Table 4: Confusion matrix.

	Recall	Precision	F1 score
Test	99.09	98.04	97.02
Train	99.9	98.89	98.02
Validate	95.06	97.12	96.45

learning-based method for detecting anomalous network intrusions. The technology is self-updating and is capable of detecting new harmful network traffic patterns.

3. Intrusion Detection System Based on Deep Reinforcement Learning

The PPO2-based intrusion detection system DRL-IDS proposed in this paper is mainly composed of three parts: the data processing module, the intrusion detection agent (agent) construction module, and the intrusion detection agent training module (see Figure 1). The processing module mainly includes feature selection and data preprocessing; the intrusion detection agent construction module mainly includes determining the environment state model of reinforcement learning, value function construction, and training strategy definition. The environment state model is the private presentation of the environment, including the domain used to determine the rules of reward and punishment that are not visible to the intrusion detection agent. The training strategy further optimizes the action decision strategy of the intrusion detection agent by evaluating the value function and uses the reward and punishment (loss function) feedback from the environmental state model to update the parameters in the training strategy. The training module of the intrusion detection agent continuously optimizes the system through the loss function until the model converges or completes the specified.

3.1. Data Processing Module. To reduce the noise redundancy of the original data and improve the multiclass detection accuracy of the model, the intrusion detection system in this paper first performs feature selection. It effectively reduces the redundant dimension of the data under the premise of ensuring intrusion detection performance. The intrusion detection system is based on embedding. The feature selection algorithm of the formula LightGBM algorithm performs feature screening [16], and the specific strategies are as follows:

(a) Delete features with missing values greater than 60% of the threshold. According to experience, when the missing rate is greater than 60%, this feature is of little significance to the training of intrusion detection agents

(b) Delete the unique value of the feature

(c) Delete any feature in each strongly correlated feature pair. The strongly correlated feature pair's specific threshold (absolute value) is defined as Pearson correlation coefficient [17] 0.99

(d) Delete the features with a lower importance ranking from the LightGBM algorithm. The final number of features will gradually increase the features in the order of feature importance scores until the model's performance no longer improves

(e) All to normalize the variables in different intervals use the most straightforward min-max function to scale the characteristic values of the range to the interval [0, 1]. The specific formula is as shown in the following equation:

$$S' = \frac{S - \min(s)}{\max(s) - \min(s)}. \quad (1)$$

Among them, S is the original value, and S' is the normalized value

(f) Generate feature vectors and finally perform one-hot encoding

3.2. Intrusion Detection Agent Structure. This section describes the components that make up the intrusion detection agent: environment state model, value function, and training strategy.

3.2.1. Environmental State Model. Use real Industrial Internet of Things data sets to simulate the network traffic environment to form the environment required for the

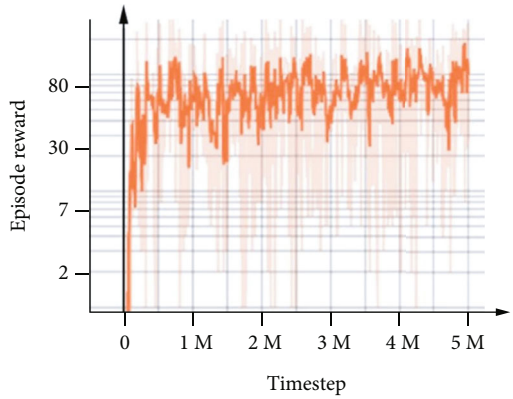

Figure 2: PPO2 episode.

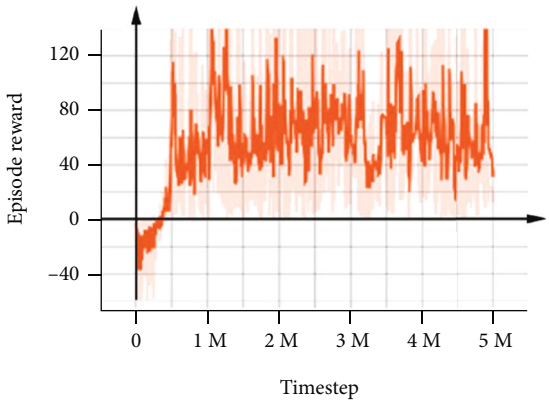

Figure 3: Episode reward.

intelligent body construction of the intrusion detection system, which includes three parts:

(a) The administrator of the auxiliary diagnosis and detection system output is mainly manifested in the feedback process of the environment agent

(b) For ordinary network users, the traffic of network users is simulated by the traffic of existing data sets

(c) Attackers create various malicious attacks

The detection agent can see the early state of the environment, according to the provided feedback signal time step t, through learning to choose a suitable action from the current time step t to the final state reward r_t, n.

The cumulative sum of $R_t = r_{t,1} + r_{t,2} + r_{t,2} + \cdots \cdots + r_{t,n}$. Because the component detection environment is random and unknown, this means that the next state may also be random, so it will happen that the action of the event detection agent has randomness; as the number of steps increases, the probability of a rabbit with a specific ability to get the same reward decreases. Decompression is used to reduce specificity and randomness to verify the strong correlation of the steps, and the future reward is replaced by the stimulus future cumulative reward C_t.

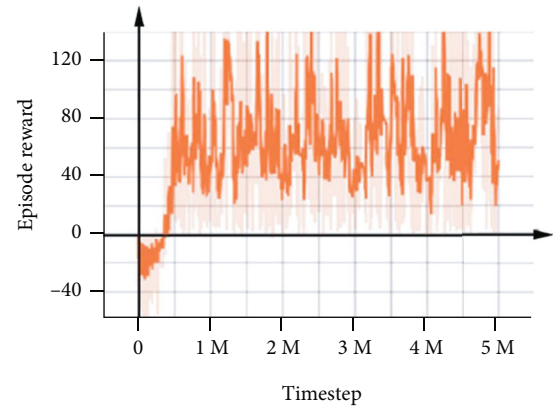

Figure 4: DQN episode.

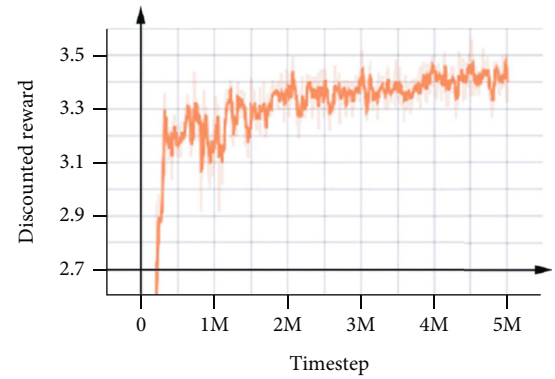

Figure 5: Discounted reward.

The expression of the total bundled future cumulative compensation of time step T is

$$C_t = R_{t+1} + \alpha R_{t+2} + \cdots + \alpha^m R_{t+m+1} = \sum_{m=0}^{\infty} \alpha^m R_{t+m+1}. \quad (2)$$

Among them, $\alpha \in [0, 1]$ is a discount coefficient used to increase instant rewards instead of delayed rewards. In addition,

(a) α is close to 0, which means more attention is paid to the current return

(b) α is close to 1, which means more attention is paid to future returns

Since the network flows measured by intrusion detection are discrete and independent from each other [18], α should be as close to 0 as possible in the experiment. As a result, the continuity between network flows is weakened. The rules for interactive feedback between the intrusion detection agent and the environment are as follows:

(a) When the intrusion detection system successfully detects an attack and successfully classifies the type of the attack, it will give a positive feedback m_{t+1}

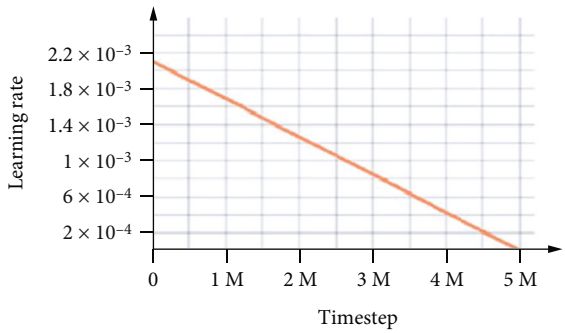

Figure 6: Learning rate.

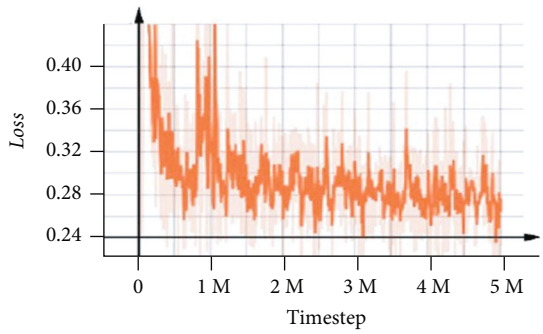

Figure 7: Loss function.

(b) When an attack is missed or an attack is successfully detected, but the type of the attack is incorrectly classified, a harmful feedback m_{t-1} is given

(c) When the flow is normal and there is no alarm, there is no feedback

This paper uses the addition and subtraction of real numbers to define the rules of rewards and punishments to achieve the purpose of intrusion detection agent training. All data that can affect the environment and generate rewards and punishments are considered part of the environment state. Thus, the DRL-IDS agent interacts with the environment. In the process, the feedback rules based on the network traffic environment constitute the environmental state model.

3.2.2. Value Function Construction. The value function is the expectation of rewards, which is mainly used to evaluate the quality of different states and guide the agent's actions. The data that can influence the agent to make the next action decision is part of the agent's state. Thus, the value function is used to evaluate intrusion detection of the agent at a specific time t and states. The intrusion detection agent uses the state value function to calculate the value of each state under the current strategy and uses the action-value process to calculate the value of different actions in each state. Select the action that maximizes the value function in the current state to perform the strategy optimization.

3.3. Training Strategy Definition. The training strategy of intrusion detection agents is the mapping from state to action. The algorithm PPO2 used by DRL-IDS was proposed by DeepMind and OpenAI. This algorithm is derived from the regional best algorithm TRPO, which is more straightforward, versatile, and complex than TRPO and PPO2. The degree of PPO2 is also low. The main contribution of PPO2 is to simplify the mathematical operation of the Kullback-Leibler penalty coefficient. As a method based on policy gradient, its characteristic is to train a random model or neural network directly. The error function of the deep reinforcement learning algorithm A3C [20] needs to be optimized after taking the logarithm of the target strategy. The error function of PPO2 is mainly used to evaluate the proportion of the new and old approach; unlike Q-learning, PPO2 does not use the experience buffer to store experience but learn from the environment only.

4. Experiments and Results

For the experimental environment, the framework of reinforcement learning used in DRL-IDS proposed in this paper is based on Stable (2.10.0) [22], which is an improved implementation of a set of reinforcement learning algorithms based on the OpenAI baseline relying on TensorFlow (1.14.0), an end-to-end machine learning open-source platform to build neural networks, using OpenAI Gym (0.17.2) library to assist in completing the customized environment in reinforcement learning, and Four indicators are used to evaluate the performance of the model, namely accuracy, precision, recall, and F1 score. For hardware, this experiment is in Ubuntu18. 04. It was completed on the machine of the 3LTS system. The hardware used in the investigation is as follows: CPU Model: Intel Xeon E5-2618L v3, GPU: NVIDIA GeForce RTX 2080 TI, RAM: 64 G.

This paper uses the real data set of the natural gas pipeline transportation network publicly released by the U.S. Department of Energy's Oak Ridge National Laboratory [23] to perform performance evaluation experiments on the proposed DRL-IDS intrusion detection model. In this data resource, collect the standard network traffic data and seven different types of attack data. The data set has 26 features and one label. In the experiment, we divided the data set into three parts, 60% used for training and 20% used for training. For all experiments in this article, the tests are performed on the same data set. The quantity of each category is listed in Table 1. At the same time, all experiments in this article are repeated to avoid errors in the experimental results as much as possible. Perform ten times and average all the calculation results to produce the final testing results.

The data set of the natural gas pipeline transportation network has a total of 26 features, and we use 3. The feature selection scheme mentioned in Section 2 removes useless features and reduces the complexity of the operation without reducing performance. The first step is to delete 0 missing ones with a rate greater than zero. The second step is to delete eight parts that have only a single value. The third step is to delete any of the four pairs of strongly correlated features. The fourth step is to rank the importance of the elements with LightGBM. Select the first 12 features for the

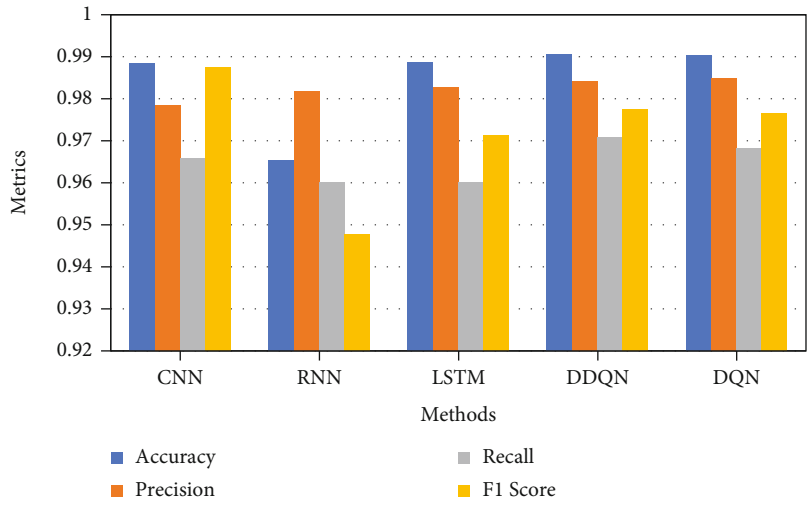

Figure 8: Comparison with other model-based IDSs.

Table 5: Performance comparison with other detection models.

	Accuracy	Precision	Recall	F1 score
CNN	0.9884	0.9784	0.9658	0.9874
RNN	0.9654	0.9817	0.9602	0.9478
LSTM	0.98874	0.9826	0.9602	0.9712
DDQN	0.9905	0.9842	0.9708	0.9774
DQN	0.9904	0.9848	0.9681	0.9765

experiment. The fifth step is to normalize the variables, generate feature vectors, and finally perform one-hot encoding. The GBM's feature selection algorithm is used to extract the most convincing set of features from Industrial Internet of Things data; after which, in combination with the deep learning algorithm, the hidden layer of the multilayer perceiving network is used as the shared core network for the value chain and strategic connection. The experiment uses the PPO2 interface of the Stable baseline to implement model training. The main parameters in the training process are listed in Table 2. Finally, the DRL-IDS intrusion detection agent is tested on the training and verification sets shown in Table 3. The results are listed in Table 4, and all indicators are above 97%.

It shows that the precision, F1 score, and recall rate for each attack detection are relatively ideal. All experiments use macro averaging to synthesize the comprehensive performance of the evaluation model, and the accuracy rate of DRL-IDS is 99.09%.

This framework is based on the PPO2 intrusion detection system DRL-IDS, which uses the TensorFlow visualization toolkit during training to track different variables. Figure 2 shows the episode reward of PPO2, and Figure 3 shows the graph between episode reward and time step. Figure 4 compares the "episode rewards" of PPO2, DQN, and DDQN, as shown in Figures 5–7. In addition, it offers the "discounted return," linear learning rate, and "loss function" of PPO2. It can be seen from the "episode reward" that compared to the other two deep reinforcement learning algorithms, intrusion detection based on the PPO2 algorithm is rewarded in the environment. At the beginning of training, it has been steadily improved. As shown from Figures 5 to 7, the intrusion detection framework based on the PPO2 algorithm converges quickly and is stable.

This intrusion detection system based on the PPO2 algorithm is compared with another algorithm DDQN [24], in the field of reinforcement learning under the same neural network structure. The parameter settings of the comparison algorithm in this paper are referred to [25-27]. It is better than other benchmark systems regarding degree, recall rate, F1 score, etc. (all experiments use a unified data set). At the same time, comparing the detection method based on deep learning and the detection method based on deep reinforcement learning, it can be seen that based on the depth, the overall performance of the reinforcement learning detection method is better. The experiment further compares the intrusion detection system DRL-IDS based on the PPO2 algorithm proposed in this paper as shown in Figures 8. The intrusion detection system based on DDQN, and the intrusion detection system based on DQN in the case of the same amount of data. The intrusion detection system based on DQN needs 18945. In 10 s, the results show that the time cost of intrusion detection system training in the Industrial Internet of Things scenario based on PPO2 is low, and it is more suitable for real intrusion detection scenarios.

5. Conclusion

In this paper, deep reinforcement learning algorithm-based intrusion detection system DRL-IDS uses a feature selection algorithm based on LightGBM, which effectively extracts the most compelling feature set in the Industrial Internet of Things data; an Industrial Internet of Things intrusion detection model is constructed based on the PPO2 algorithm. On the real data set of the Industrial Internet of Things publicly released by the U.S. Department of Energy's Oak Ridge National Laboratory, the results of a large number of experiments have shown that the intrusion detection system DRL GDS proposed in this paper performs well in

detecting various types of network attacks on the Industrial Internet of Things. Compared with the existing intrusion detection system based on deep learning or deep reinforcement learning, it is better in terms of accuracy, precision, recall rate, and F1 score as shown in Table 5. It significantly reduces the training time of intrusion detection models. In future work, we will explore Industrial IoT intrusion detection systems based on a distributed architecture.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Funding

This research work is self-funded.

References

[1] K. R. Choo, S. Gritzalis, and J. H. Park, "Cryptographic solutions for Industrial Internet-of-Things: research challenges and opportunities," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3567–3569, 2018.

[2] Q. Liu, L. Cheng, T. Ozcelebi, J. Murphy, and J. Lukkien, "Deep reinforcement learning for IoT network dynamic clustering in edge computing," in *2019 19th IEEE/ACM international symposium on cluster, Cloud and Grid Computing (CCGRID)*, pp. 600–603, Larnaca, Cyprus, May 2019.

[3] Y. Xiao, G. Niu, L. Xiao, Y. Ding, S. Liu, and Y. Fan, "Reinforcement learning based energy-efficient internet-of-things video transmission," *Intelligent and Converged Networks*, vol. 1, no. 3, pp. 258–270, 2020.

[4] E. Rabieinejad, S. Mohammadi, and M. Yadegari, "Provision of a recommender model for blockchain-based IoT with deep reinforcement learning," in *2021 5th International Conference on Internet of Things and Applications*, pp. 1–8, Isfahan, Iran, May 2021.

[5] A. Ashiquzzaman, H. Lee, T. Um, and J. Kim, "Energy-efficient IoT sensor calibration with deep reinforcement learning," *IEEE Access*, vol. 8, pp. 97045–97055, 2020.

[6] A. Borkar, A. Donode, and A. Kumari, "A survey on intrusion detection system (IDS) and internal intrusion detection and protection system (IIDPS)," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*, pp. 949–953, Coimbatore, India, Nov 2017.

[7] Y. Jiang, W. Wang, and C. Zhao, "A machine vision-based realtime anomaly detection method for industrial products using deep learning," in *2019 Chinese Automation Congress (CAC)*, pp. 4842–4847, Hangzhou, China, Nov 2019.

[8] M. A. Istiake Sunny, M. M. S. Maswood, and A. G. Alharbi, "Deep learning-based stock price prediction using LSTM and bi-directional LSTM model," in *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, pp. 87–92, Giza, Egypt, Oct 2020.

[9] L. Chen, X. Kuang, A. Xu, S. Suo, and Y. Yang, "A novel network intrusion detection system based on CNN," in *2020 Eighth International Conference on Advanced Cloud and Big Data (CBD)*, pp. 243–247, Taiyuan, China, Dec 2020.

[10] Y. S. Patil and S. V. Sankpal, "EGSP: enhanced grid sensor placement algorithm for energy theft detection in smart grids," in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, pp. 1–5, Bombay, India, March 2019.

[11] X. Yuan, J. Chen, N. Zhang, X. Fang, and D. Liu, "A federated bidirectional connection broad learning scheme for secure data sharing in Internet of Vehicles," *China Communications*, vol. 18, no. 7, pp. 117–133, 2021.

[12] E. Suwannalai and C. Polprasert, "Network intrusion detection systems using adversarial reinforcement learning with deep Q-network," in *2020 18th International Conference on ICT and Knowledge Engineering (ICT&KE)*, pp. 1–7, Bangkok, Thailand, Nov 2020.

[13] W. Zhou, J. Li, Y. Chen, and L. Shen, "Strategic interaction multi-agent deep reinforcement learning," *IEEE Access*, vol. 8, pp. 119000–119009, 2020.

[14] B. Wu and Y. Feng, "Policy reuse for learning and planning in partially observable Markov decision processes," in *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, pp. 549–552, Changsha, China, July 2017.

[15] D. Wang, B. Ding, and D. Feng, "Meta reinforcement learning with generative adversarial reward from expert knowledge," in *2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pp. 1–7, Dalian, China, Sept 2020.

[16] V. Shakya and R. R. S. Makwana, "Feature selection based intrusion detection system using the combination of DBSCAN, K-mean++ and SMO algorithms," in *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, pp. 928–932, Tirunelveli, India, May 2017.

[17] W. Teng, L. Cheng, and K. Zhao, "Application of kernel principal component and Pearson correlation coefficient in prediction of mine pressure failure," in *2017 Chinese Automation Congress (CAC)*, pp. 5704–5708, Jinan, China, Oct 2017.

[18] E. Anthi, L. Williams, and P. Burnap, "Pulse: an adaptive intrusion detection for the Internet of Things," in *Living in the Internet of Things: Cybersecurity of the IoT*, pp. 1–4, Location: London, UK, March 2018.